A Human-Centered Evaluation of a Deep Learning System Deployed in Clinics for the Detection of Diabetic Retinopathy

Abstract

Deep learning algorithms promise to improve clinician workflows and patient outcomes. However, these gains have yet to be fully demonstrated in real world clinical settings. In this paper, we describe a human-centered study of a deep learning system used in clinics for the detection of diabetic eye disease. From interviews and observation across eleven clinics in Thailand, we characterize current eye-screening workflows, user expectations for an AI-assisted screening process, and post-deployment experiences. Our findings indicate that several socio-environmental factors impact model performance, nursing workflows, and the patient experience. We draw on these findings to reflect on the value of conducting human-centered evaluative research alongside prospective evaluations of model accuracy.